PCT/EP200 4 / ...

The Patent Office

REC'D 27 APR 2004

WIPO PCT

INVESTOR IN PEOPLE

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

15. 03. 2004

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated 3 March 2004

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

1/77

26MAR03 E795087-1 D00085
P01/7700 0.00-0306855.8

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

THE PATENT OFFICE
D
25 MAR 2003
LONDON

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

JCB/935

2. Patent application number
(The Patent Office will fill in this part)

0306855.8

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

Ideas Network Ltd.,
42 Horseguards Drive,
Maidenhead,
Berks,
England, SL6 1XL

Patents ADP number *(if you know it)*

If the applicant is a corporate body, give the country/state of its incorporation

UNITED KINGDOM

8579418001

4. Title of the invention

DATA COMMUNICATION NETWORK

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

Stevens, Hewlett & Perkins
Halton House,
20/23 Holborn,
London, EC1N 2JD

Patents ADP number *(if you know it)*

1545003

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*
a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

YES

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	16
Claim(s)	8
Abstract	
Drawing(s)	19 + 19

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature Stevens, Hewlett & Perkins Date 25.03.03

Stevens, Hewlett & Perkins

12. Name and daytime telephone number of person to contact in the United Kingdom

J.C. BOYDELL 020 7404 1955

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) *If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.*

b) *Write your answers in capital letters using black ink or you may type them.*

c) *If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*

d) *If you have answered 'Yes' Patents Form 7/77 will need to be filed.*

e) *Once you have filled in the form you must remember to sign and date it.*

f) *For details of the fee and ways to pay please contact the Patent Office.*

"DATA COMMUNICATION NETWORK"

This invention relates to a data communication network.

A simple data communication network is illustrated in Figure 1 of the accompanying drawings. The network comprises a plurality of nodes N1 to N10 interconnected by transmission lines T1 to T13, normally referred to simply as links. The nodes will generally comprise some form of intelligent switching device. The transmission lines may comprise wire links, fibre optic links, infra red or wireless links or combinations of these. The particular nature of the transmission lines is not important: they are simply a means of enabling data to be passed between the nodes. As illustrated in Figure 2, most of the nodes and their associated transmission lines form closed rings 11 to 15. The nodes N1, N2 and N3 and their associated transmission lines T1 and T2 are connected in a tree structure 10 with node N1 at the remote end of the tree. Routing data across the tree structure 10 is deterministic as there is only one path to the destination that the data can follow. The invention, allows for the rest of the network to be reduced to a logical tree structure, through a series of abstractions, so that together with the physical tree structure 10 there can be deterministic routing across the whole network.

Any of the nodes N1 to N10 can be used to connect to a host machine or machines (not shown) possibly themselves interconnected by means of a LAN (local area network).

As is well known, in such networks data to be transmitted is segmented, and each segment is encapsulated in a packet, frame or cell (depending on the protocol being used) for transfer across the network. Each packet, frame or cell will contain, as well as the data to be transmitted, control data which will generally comprise such information as the source and destination address of the packet. Various different methods are used for routing the packets, frames or cells across the network. Generally speaking the switching device at each node includes a storage device which stores a routing table which is accessed according to

the control data contained within the packet, frame or cell in order to forward the packet, frame or cell to its destination by the preferred route. The manner in which the table is populated and acted upon gives rise to the different methods of routing. The software associated with populating these tables is controlled by a routing algorithm. The primary purpose of the routing algorithm, of which there are several in common use, is to determine the preferred route to the destination, based on one or more parameters, and to populate the routing table in each node to give effect to the result of this determination. Then, when a packet, frame or cell arrives at a particular node, the contents of the table will dictate the path upon which it is transmitted from the node. Although the present invention is connected with routing, it is in fact transparent to the particular method of switching which is used.

Thus it will be seen that each node will contain routing information which will be used to forward the packet, frame or cell to its destination. The type of information stored will depend upon the sophistication of the network, and the particular routing method which is used. The routing information in each node will also typically take into account the performance of the network – for example whether there are any congestion problems or link failures ahead. In order to keep the routing information at each node up-to-date, it is normally provided that the information is updated periodically to cater for changes in the network, through packets, frames or cells that are purely control messages.

The present invention is directed to what is considered to be a new way of looking at networks, by basing the routing algorithms, which in turn control the content of the information in the routing tables, on a recursive abstraction of the physical network into a series of logical levels

According to a first aspect of the present invention there is provided a data communication network comprising a plurality of nodes and transmission lines linking said nodes, each node including storage means holding routing information, said routing information being created by defining the network as a series of recursive abstractions forming one or

more logical levels, the bottom level of which corresponds to the physical network, said logical levels being formed by grouping the nodes of each level of the network into closed rings, and considering each such closed ring as a logical node of the next higher logical level.

According to a second aspect of the present invention, there is provided a node intended to form part of a data communication network, said node comprising a plurality of input/output interfaces for data entering or leaving the node, storage means for storing a routing table containing routing information for data input to the node, processor means running software to populate said routing table, said software being based on algorithms which are based on a definition of the network as a series of recursive abstractions into one or more logical levels, the bottom level of which corresponds to the physical network, and in which said logical levels are formed by grouping the nodes of each level of the network into closed rings and considering each such closed ring as a logical node of the next higher logical level.

Within each closed ring, each node exchanges address and performance information with the other node or nodes in the closed ring. However, nodes are not necessarily aware of information relating to other rings, and it is this that leads to a definition of two types of relationship between rings. If all nodes in two rings are aware of the address and performance information in both rings, then the two rings are said to have a flat relationship with respect to each other, and are said to share information symmetrically. Many rings may have such a relationship with one another, making a flat network. By contrast, if two rings have a hierarchical relationship to one another, then one ring will be defined as being higher in the hierarchy than the other, and the nodes in the higher level ring will receive information from the nodes in the lower level ring, but the nodes, other than the common nodes, in the lower level ring will not receive information from the nodes in the higher level ring. Two rings in a hierarchical relationship with respect to each other are said to share information asymmetrically. In both flat and hierarchical relationships the

nodes that distribute information between the closed rings are referred to as gateways. The process of distributing address and performance information across the network is referred to as advertising and is described in detail below.

The invention, unlike other methods in common use which seek to determine a unique path based on optimising a particular set of one or more parameters, is able to retain all the routing possibilities and to exploit them at each logical level, as appropriate, based on the performance information available. This routing decision is made in accordance with the routing information stored at each node and this routing information is created by running a routing algorithm which is based on a recursive abstraction of the network into logical levels, as described above. When a packet, frame or cell arrives at a node, the node is expected to forward the packet, frame or cell onwards through an appropriate output port in accordance with the information contained in the routing table. This process is repeated at each node as the packet, frame or cell travels across the network. Although in reality the packets, frames or cells travel across the physical network, it is possible, for the purpose of explaining the present invention, to conceptually consider the packet, frame or cell at the same time also travelling up and down the logical levels of the network in order to control its physical passage across the network. The overall process of successively forwarding data node-by-node across the network is referred to as routing and is described in detail below.

The concepts described herein are independent of the particular method of switching which is used, and of the particular method of addressing which is used. Examples of particular techniques for advertising and routing on both flat and hierarchical network architectures, as well as a number of other issues associated with the correct functioning of the network are described below.

In order that the invention may be better understood, several embodiments thereof will now be described by way of example only and with reference to the accompanying drawings in which:-

Figure 1 is a diagrammatic view of a data communication network;

Figure 2 is a diagram to show how the network of Figure 1 may be divided into closed rings;

Figure 3 is a block diagram of a node suitable for use in the network of the invention;

Figures 4 to 23 are diagrams for explaining how a representative network may be abstracted into four logical levels, according to the methods described in this invention, in order to form a loop-free structure that will enable deterministic routing;

Figures 24 and 25 are diagrams showing how the network of Figure 1 may be recursively abstracted into three logical levels, according to the methods described in this invention, in order to form a loop-free structure that will enable deterministic routing.

Referring firstly to Figure 3, there is shown a block diagram of a node 30 suitable for use in the network of the present invention. The node 30 comprises 3 input/output ports 31,32,33 each of which connect via a respective queuing memory 34,35,36 to a switching means in the form of a switching fabric 37. Although shown with 3 input/output ports, it will be understood that the node may in practice contain any number of ports, from two upwards. Each port is connected via a transmission line to an adjacent node in the ring, or may be connected by a transmission line to a host which may be a stand-alone computer terminal, or a LAN.

The switching fabric 37 receives packets, frames or cells from one of ports 31,32 or 33 and forwards them to another of the ports 31,32 or 33 for onwards transmission The memories 34,35 and 36 provide a queuing facility in the event that the node receives more data within a particular time frame than it can handle.

The decision as to which port the incoming data is forwarded by the switching fabric 37 is taken on the basis of routing information contained in a routing table 38 which is maintained by a microprocessor and associated circuitry 39. The microprocessor 39 is linked to the table 38 by an input/output interface 40. Associated with the microprocessor 39 is a

memory 41 which contains the software which controls the microprocessor, and may also contain a copy of the current contents of the routing table. The software is based on intelligent algorithms which essentially comprise a set of rules which dictate how any particular packet, frame or cell received by the node is to be forwarded, bearing in mind such factors as its required destination address, and the state of the network.

All packets, frames or cells are examined on arrival to ascertain whether they are control messages which contain purely control information – for example, changes in address and performance information which needs to be reflected in the routing table. Any such packets, frames or cells are forwarded by the switching fabric 37 to the microprocessor 39 via a queuing memory 42 and input/output interface 43. These control messages are used to set up the node in the first place, and to keep it up to date with any changes in the network which require changes in the routing table, or indeed in the algorithms themselves.

Reference is now made to Figures 4 to 23 which explain the manner in which a physical network may be considered as a hierarchy of logical levels and consider the processes of advertising and routing in connection with this recursive abstraction of the network into a series of logical levels..

Before discussing a typical real life network example, such as that shown in Figure 1, it is easier to understand the overall principles of the routing technology of the invention by considering a more theoretical network of 12 nodes as shown in Figure 4.

It will now be explained how the physical network shown in Figure 4 can be recursively abstracted into a series of logical levels for the purpose of routing data across the network.

The lower-most logical level, which we refer to for convenience as level 1, is identical to the physical network. The appearance of the lower-most logical level is thus the same as the physical network. It will be seen from Figure 4 that the topology of the network can be decomposed at level 1 into 5 closed logical rings R1 to R5, with many of the nodes and links in the network forming part of more than one ring.

We next consider the advertising of address and performance information from a node in level 1. As an example, take the ring R1 of level 1 of the network, as shown in Figure 4. Using node N1.1 as an example, Figure 5, shows how advertisements from a node are circulated around logical ring R1 to the other nodes in that ring. The advertisements are identified as originating from node N1.1 by a source_id set equal to N1.1.

Each node on the logical ring R1 advertises any hosts (not shown) attached to it by sending announcement control messages around the ring R1. The messages are associated with ring R1 by a ring_id set equal to R1 and are themselves addressed to their neighbours using a generic neighbour address that all nodes will recognise. As a general rule, data cannot be returned over the interface it arrived on (unless it has changed to a different logical ring). As a result, the announcement messages are automatically circulated around the logical ring, maintaining their direction of rotation. To prevent infinite looping of the announcement messages they are tagged with a loop_avoidance_id set equal to the originating node_id (N1.1 in the example in Figure 5) – so that they can be recognised upon their return to the originating node and dropped. Figure 5 only shows announcement messages being sent in a clockwise rotation, but it is recommended that announcement messages are sent both ways around the ring, so that the system is able to withstand single link or node failures.

We now consider the next higher logical level, referred to as level 2, by reference to Figure 6 which shows the nodes associated with the level 1 rings R1 to R5 in grey. Since all of the nodes in each of the logical rings R1 to R4 fully share their address and performance information they can each be represented as a single logical node in the next higher logical level (level 2). Thus, Figure 6 shows the level 2 abstraction in bold: for example, the level 1 nodes associated with ring R1 are now represented by logical node N2.1. The resulting 5 logical nodes in level 2 can now themselves be formed into closed logical rings. For example, nodes N2.1, N2.2 and N2.5 now form the level 2 ring R125. The resulting 4 rings provide the level 2 abstraction of the network. Thus it can be seen that the

architecture is recursive, because now the same processes are applied again (see below) – although they must be developed further to allow for the logical nature of the network abstraction. Indeed, its abstraction from the physical topology of the network creates added flexibility that can now be exploited to create either flat or hierarchical relationships between the levels. First we will consider a level 2 abstraction based on a flat relationship and then we will consider the impact on the level 2 abstraction which results when a hierarchical relationship is imposed.

Figure 7 uses ring R125 as an example of how advertising of address and performance information is achieved at level 2. As, in this example, the level 1 rings are in a flat relationship there are symmetrical exchanges of information between them – corresponding to the flow of advertisements between the level 2 logical nodes The example shows advertisements originating from level 1 ring R1 (represented by node N2.1) and circulating around ring R125 in a clockwise direction of rotation. As with the level 1 rings, it is recommended that a more robust solution is achieved by circulating the advertisements in both directions of rotation – clockwise (shown) and anti-clockwise (not shown).

At logical level 2, the link between the level 2 nodes is by way of shared nodes in level 1. To this end, the shared nodes act as gateways, and it will be seen that, as a general matter, gateways are used to circulate advertisements around the higher-level logical rings. The existence of gateways on a ring are advertised in the same way as other destination addresses. Once a particular gateway has passed the information into a neighbouring level 1 ring, the normal level 1 advertisement mechanism, as described above, is used to distribute the information through that neighbouring ring. As a matter of course this will distribute the information to those nodes in the ring that are acting as the level 2 gateway to the next level 1 neighbouring ring, and so the process repeats itself. It will be realised that steps must be taken to prevent infinite loops of information now at levels 2 and 1.

Figure 8 shows the level 2 flow of advertisements from ring R1 into ring R2 using gateways in nodes N1.2 and N1.10. A single gateway can be used, but two or more gateways provide a more robust solution that can withstand a single link or node failure. Figure 8 shows the gateways then circulating the level 2 information around the level 1 ring R2 in a clockwise direction of rotation. Again, it is recommended that a more robust solution is achieved by circulating the advertisements in both directions of rotation – clockwise (shown) and anti-clockwise (not shown). The advertisements are identified as originating from nodes N1.2 or N1.10 by their source_ids. As the advertisements belong to level 2, they have an inner ring_id of R125, and as they are being circulated around the level 1 ring R2, they have an outer ring_id of R2. In both cases, there must be a loop_avoidance_id to prevent infinite looping at either level 1 or level 2. By using a loop_avoidance_id which is common to both originating gateways, it will provide an automatic drop and substitute action – minimising the flow of address advertisements to a single pair of rotations (as they correctly drop their partner gateway advertisements assuming that they are their own). As the level 2 advertisements in this example originated in level 1 ring R1, the level 2 loop_avoidance_id is set to ring R1 so that the information can be dropped upon its return to ring R1. As the level 1 loop_avoidance_id must be common to the two gateways to achieve the correct drop and substitute action, it is also set to ring R1 as this is a convenient common point of reference for these two gateways.

Figure 9 shows the level 2 flow of advertisements from ring R2 into ring R5 using gateways in nodes N1.10 and N1.11. Figure 9 shows the gateways circulating the level 2 information around the level 1 ring R5 in an anti-clockwise direction of rotation. Again, it is recommended that a more robust solution is achieved by circulating the advertisements in both directions of rotation – clockwise (not shown) and anticlockwise (shown). The advertisements are identified as originating from nodes N1.10 or N1.11 by their source_ids. As these advertisements belong to level 2, they continue to have an inner ring_id of R125, and as they are now being

circulated around the level 1 ring R5, they have an outer ring_id of R5. As these advertisements originated in the level 1 ring R1, they continue to have a level 2 loop_avoidance tag of R1. As the level 1 loop_avoidance tag must be common to the two gateways to achieve the correct drop and substitute action, it is set to R2 as this is a covenient common point of reference for these two gateways.

Figure 10 shows the level 2 advertisements being dropped on their return from ring R5 to ring R1 by the gateways in nodes N1.9 and N1.10. These gateways determine from the level 2 loop_avoidance_id that this level 2 information originated from R1 and so must be dropped on its return.

We now consider the next higher logical level, referred to as level 3, by reference to Figure 11 which shows, in grey, the level 1 and level 2 nodes. Since all the nodes in each of the level 2 rings fully shares their address and perfomance information (see above) they can each be represented as a single logical node in the next higher logical level (level 3). Thus Figure 11 shows the level 3 abstraction in bold. For example, the level 2 nodes associated with ring R125 are now represented by the logical node N3.1. The resulting 4 logical nodes in level 3 can now themselves be formed into a single closed logical ring. The resulting ring R12345 provides the level 3 abstraction of the network. Thus it can be seen that the architecture is recursive, because now the same processes are applied again (see below).

Figure 12 shows how advertising of address and performance information is achieved at level 3 around ring R12345. As in this example, the level 2 rings are in a flat relationship there are symmetrical exchanges of information between them corresponding to the flow of advertisements between the level 3 logical nodes.

The advertisements are shown originating from level 2 ring R125 (represented by node N3.1) and circulating around ring R12345 in a clockwise direction of rotation. As with the level 1 and level 2 rings, it is recommended that a more robust solution is achieved by circulating the

advertisements in both directions of rotation – clockwise (shown) and anticlockwise (not shown).

At logical level 3, the link between the level 3 nodes must still be by way of shared nodes in level 1. Once a particular gateway has passed the information into its corresponding level 2 ring, the normal level 2 and level 1 advertising mechanisms, as described above, are used to distribute the information through that neighbouring ring. As a matter of course this will distribute the information to those nodes in the ring that are acting as the level 3 gateway to the next level 2 neighbouring ring, and so the process repeats itself. It will be realised that steps must now be taken to prevent infinite loops of information at levels 3, 2 and 1.

Figure 13 shows the level 3 flow of advertisements from ring R125 into ring R235 using gateways in node N1.2 and N1.10, and the subsequent distribution of the level 3 information to the nodes representing ring R235 by its circulation around level 1 ring R2. The advertisements are identified as originating from nodes N1.2 or N1.10 by their source_ids. The inner ring_id is therefore R12345 with a level 3 loop_avoidance_id of R125. the outer ring_id is therefore R2 and the level 1 loop_avoidance_id is R1. (Note that in this topology there is no need for an additional level 2 distribution of the level 3 information).

Figure 14 shows the level 3 flow of advertisements from ring R235 into ring R345 using gateways in node N1.4 and N1.11, and the subsequent distribution of the level 3 information to the nodes representing ring R345 by its circulation around level 1 ring R3. The advertisements are identified as originating from nodes N1.4 or N1.11 by their source_ids. The inner ring_id is therefore R12345 with a level 3 loop_avoidance_id of R125. the outer ring_id is therefore R3 and the level 1 loop_avoidance_id is R2.

Figure 15 shows the level 3 flow of advertisements from ring R345 into ring R145 using gateways in node N1.6 and N1.12, and the subsequent distribution of the level 3 information to the nodes representing ring R145 by its circulation around level 1 ring R4. The advertisements are identified as originating from nodes N1.6 or N1.12 by their source_ids. The

inner ring_id is therefore R12345 with a level 3 loop_avoidance_id of R125. the outer ring_id is therefore R4 and the level 1 loop_avoidance_id is R3.

Figure 16 shows the level 3 advertisements being dropped on their return from ring R145 to ring R125 by the gateways in nodes N1.8 and N1.9. these gateways determine from the level 3 loop_avoidance_id that this level 3 information originated from R125 and so must be dropped upon its return.

We now consider the next higher logical level, referred to as level 4, by reference to Figure 17 which shows in grey the level 1,2 and 3 nodes. Since all the nodes in the level 3 ring fully share their address and peformance information (see above), they can be represented as a single logical node in the next higher logical level (level 4).

Thus Figure 17 shows the level 4 abstraction in bold as a single logical node N4.1. As level 4 forms a loop-free higher-level logical topology, the recursive abstractions of the network are complete. There is now deterministic routing across the whole network. For further clarity the stages of the recursive abstraction of the network into a series of logical levels are summarised in Figure 18.

Reference is now made to Figure 19 which illustrates an example of data being routed across the network following completion of the advertising process described above. The source of the data is a host machine attached to node N1.2 and the destination is a host machine with address xyz attached to node N1.6. Node 1.2 is aware of destination xyz from advertisements received at level 3 which originated from level 2 ring R345. These level 3 advertisements would have been distributed to node N1.2 with source_ids corresponding to the level 3 gateway in nodes N1.8 & N1.9 and N1.4 & N1.11. (I.e. from rotations of level 3 advertisements both ways around the level 3 ring). Node N1.2 is aware of how to reach these gateway nodes from their advertisements on the level 1 rings R1 and R2 that node N1.2 is connected to. Based on performance information, node N1.2 decides to route the data to the level 3 gateway in node N1.9, by circulating it clockwise around level 1 ring R1.

When the data gets to node N1.10, it makes its own independent routing decision, but is constrained by the general rule that data cannot be returned over the interface it arrived on (unless it has changed to a different logical ring). Based on performance information node N1.10 decides to continue to route the data to the level 3 gateway in node N1.9 and so maintains the direction of circulation on ring R1. (Note, however, that node N1.10 could have decided to direct the data to the alternative level 3 gateway N1.11 if it had preferred).

When the data gets to node N1.9, it makes its own independent routing decision, and as it is also on level 1 ring R4, it knows it can reach destination xyz through node N1.6 on ring R4. Based on performance information, node N1.9 decides to route the data to node N1.6 by circulating it clockwise around ring R4. (Note that it could also have circulated it anticlockwise around R4 to the same destination if it had preferred).

When the data gets to node N1.12, it makes its own independent routing decision, but is constrained by the general rule that data cannot be returned over the interface it arrived on (unless it has changed to a different logical ring). Node N1.12 is therefore obliged to maintain the direction of rotation to N1.6 on ring R4.

When the data arrives at node N1.6, it recognises that it is the address of a locally attached host and delivers the data.

Reference is now made to Figure 20 to see how the situation is simplified by imposing a hierarchical relationship between ring R5 as the higher level and rings R1, R2, R3 and R4 as the lower level. Each of the level 1 rings can still be replaced by a single logical node in level 2. For example, level 1 ring R1 is still replaced by logical level 2 node N2.1. However, the hierarchical relationship means that information between the lower level rings must be via the higher level ring R5. As a result the level 2 abstraction of the network becomes immediately a loop-free logical topology and so no further recursive abstraction of the network is necessary. Routing is already deterministic across the whole network. The

imposition of hierarchy, can therefore simplify the issue of advertising across the network, although as a result there will be less route flexibility. In many practical network situations this may not be a significant constraint and indeed may be a benefit.

Reference is now made to Figures 21, 22 and 23 to see how the imposition of the hierarchy impacts the advertising and routing processes. Figure 21 uses the interaction between ring R1 and ring R5 as an example of how advertising of address and performance information is impacted at level 2 by the imposition of a hierarchy. First it is important to note that the information distribution is now asymmetric and advertisements only flow upwards from R1 to R5. There are no advertisements in the opposite direction from R5 to R1.

At logical level 2, the link between the level 2 nodes is still by way of shared nodes in level 1 that act as gateways. The existence of gateways on a ring are advertised on that ring in the same way as other destination addresses. Once a particular gateway has passed the information from the lower level ring to the higher level ring, the normal level 1 advertisement mechanism is used to distribute the information through the higher level ring. Figure 22 shows as an example the level 2 flow of advertisements from ring R1 up into ring R5 using gateways N1.9 and N1.10. As the imposed hierarchy constrains the flow of level 2 information to ring R5, there is no need for the identification and management of a separate logical level 2. The level 2 information can therefore be circulated with a ring_id of R5 and an appropriate loop_avoidance_id which in this example is R1. The advertisements are identified as originating from nodes N1.9 or N1.10 by their source_ids.

Figure 23 illustrates an example of data being routed across the network following the completion of the hierarchical advertising process described above. The source of the data is a host machine attached to node N1.2 and the destination is a host machine with address xyz attached to node N1.6. Node 1.2 is not aware of destination address xyz as it is not a part of its logical level 1 rings R1 or R2 that it is connected to. Node N1.2

knows it must therefore route the data to a gateway to a higher level in order to find the destination. From advertisements on ring R1 it is aware of gateways in nodes N1.9 and N1.10, and from advertisements on ring R2 it is aware of gateways in nodes N1.10 and N1.11. Based on performance information, node N1.2 decides to route the data to the level 2 gateway in node N1.10, by circulating it clockwise around level 1 ring R1.

When the data gets to node N1.10, it makes its own independent routing decision, and can now benefit from all the address and performance information available at level 2. From this information it is aware that destination address xyz is reachable through node N1.6 via gateways in nodes N1.9, N1.11 and N1.12 – as destination address xyz is attached to node N1.6 which is part of both rings R3 and R4. Based on performance information, node N1.10 decides to route the data to the level 2 gateway in node N1.9 and so routes the data anti-clockwise on ring R5. (Note that it could also have circulated it clockwise around R5 to the level 2 gateway in node N1.11)

When the data gets to node N1.9, it makes its own independent routing decision, and as it is also on level 1 ring R4, it knows it can reach destination address xyz through node N1.6 on ring R4. Based on performance information, node N1.9 decides to route the data to node N1.6 by circulating it clockwise around level 1 ring R4. (Note that it could also have circulated it anticlockwise around R4 to the same destination if it had preferred).

When the data gets to node N1.12, it makes its own independent routing decision, but is constrained by the general rule that data cannot be returned over the interface it arrived on (unless it has changed to a different logical ring). Node N1.12 is therefore obliged to maintain the direction of rotation to N1.6 on level 1 ring R4.

When the data arrives at node N1.6, it recognises that it is the address of a locally attached host and delivers the data.

Reference is now made to Figures 24 and 25 which illustrate how the invention is applied to the more general data communication network

shown in Figure 1, and how it can be recursively abstracted into a series of logical levels. As already mentioned, the network shown in Figure 1 can be decomposed into a number of level 1 closed rings 11-15, as shown in Figure 2. Each such closed ring may now be defined as a node in a higher logical level 2 of the network, shown in Figure 24. Thus, in Figure 24, logical node N20 corresponds to ring 11, logical node N21 corresponds to ring 12, logical node N22 corresponds to ring 13, logical node N23 corresponds to ring 14 and logical node N24 corresponds to ring 15.

A hierarchical relationship is imposed between rings 11, 12 & 13 in the higher level, and tree 10, ring 14 and ring 15 in the lower level. As a result the logical nodes N21, N23 and N24 now form a loop-free higher-level logical topology.

The three logical nodes N20, N21 and N22 in Figure 24 themselves form a higher level closed ring 11-12-13. Figure 25 shows how the ring 11-12-13 can itself be defined as an equivalent node N30 in a next higher logical level 3 of the network. The node N30 now forms a loop-free higher-level topology.

Thus, at the highest level shown in Figure25, the whole route through the network is now in the form of a loop-free higher-level topology (comprising of logical and physical parts) that provides deterministic routing across the network.

It will be clear from the above explanations that any network which can be decomposed into closed rings can be treated in this way, by using recursion to create ever-higher levels of abstraction, until a loop-free higher-level topology is achieved. This loop-free topology is referred to as a tree in standard discrete mathematics terminology.

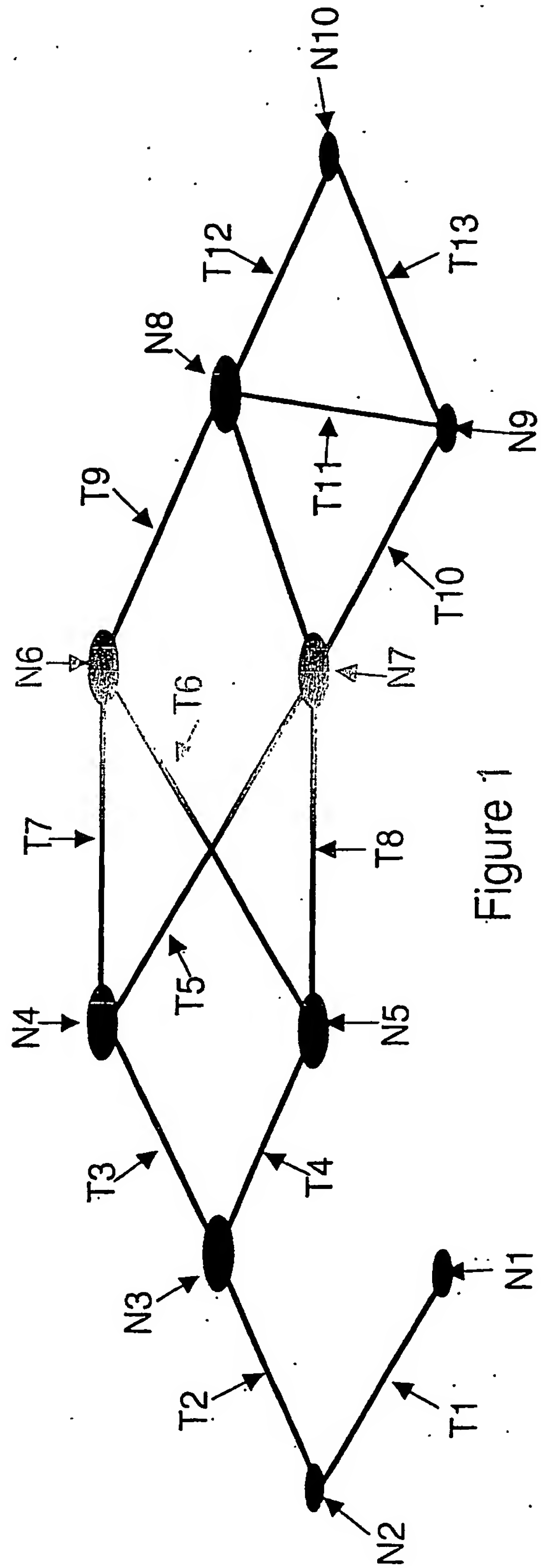


Figure 1

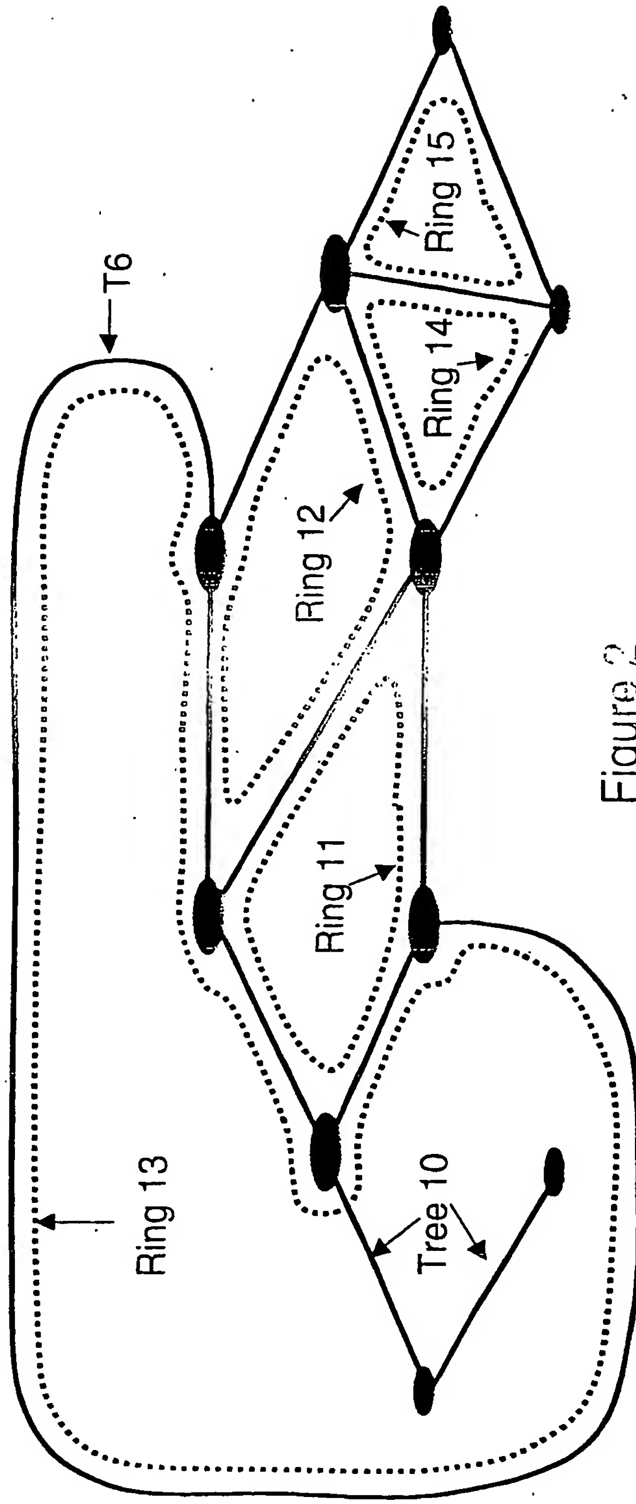


Figure 2

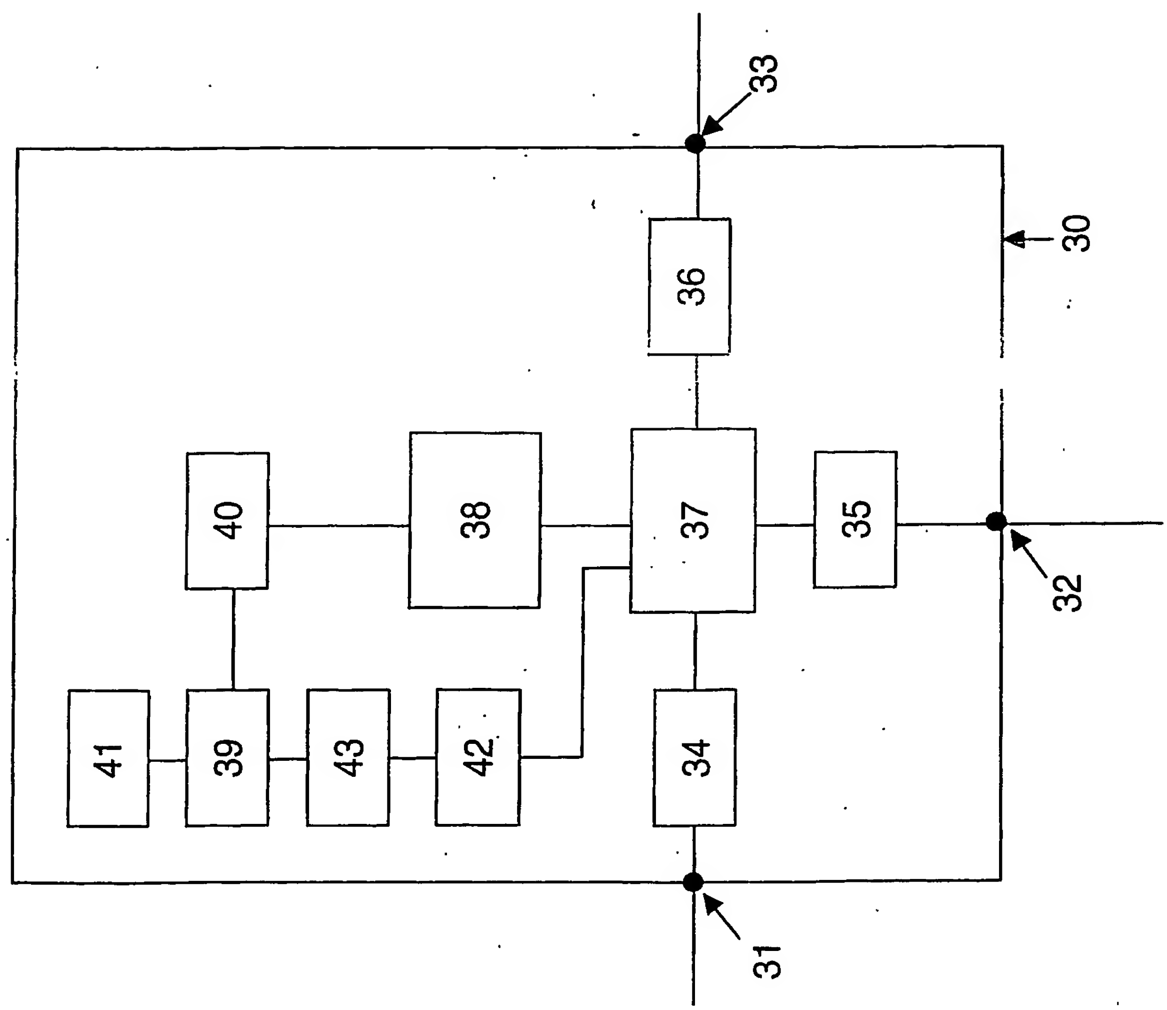


Figure 3

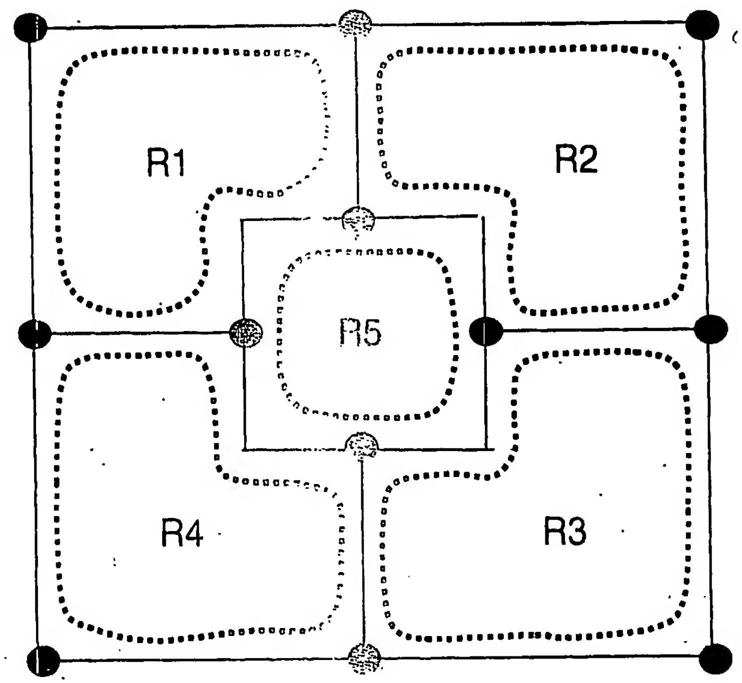


Figure 4

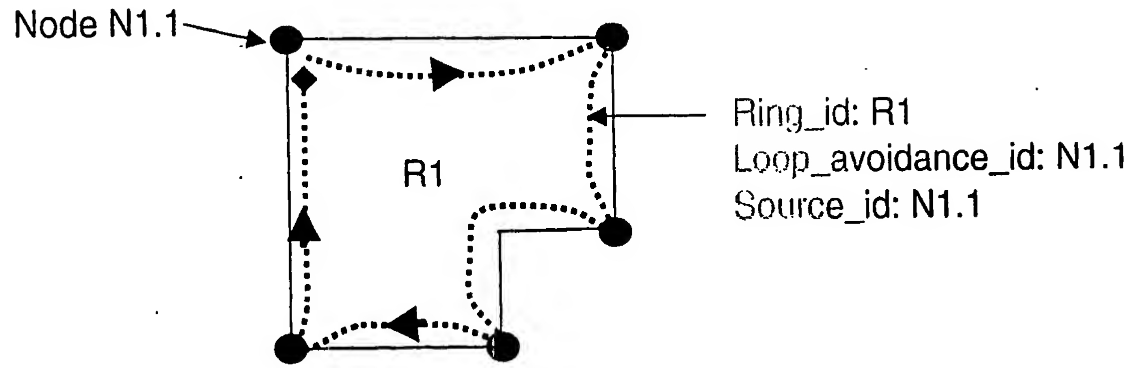


Figure 5

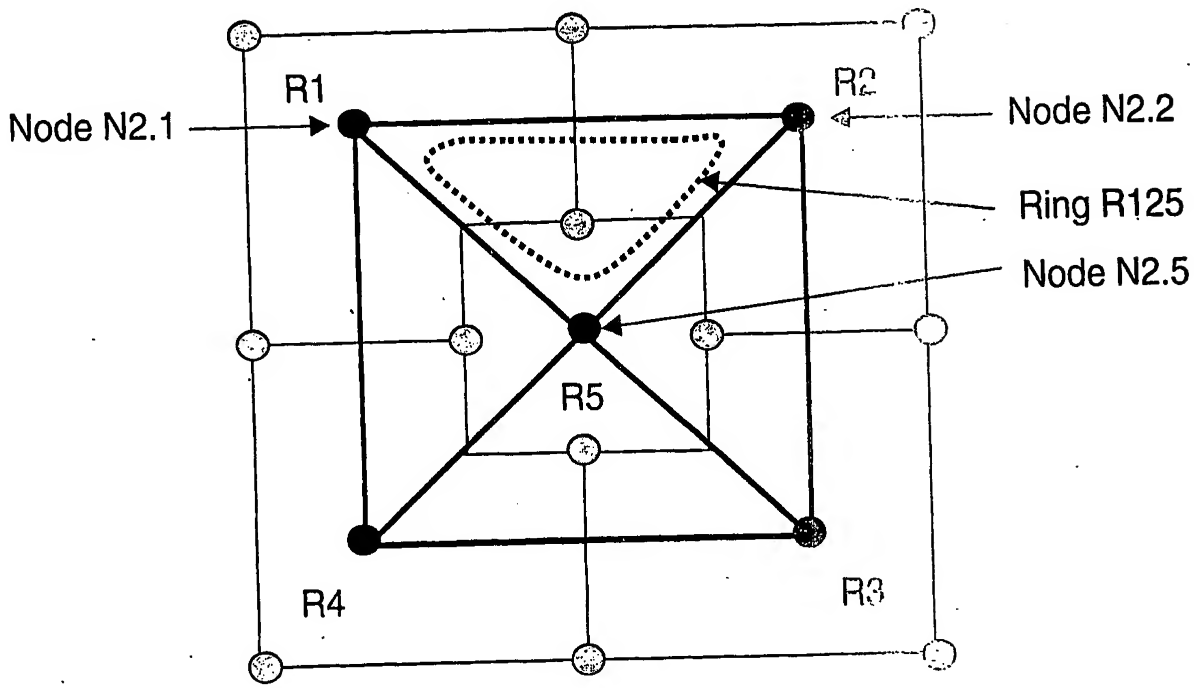


Figure 6

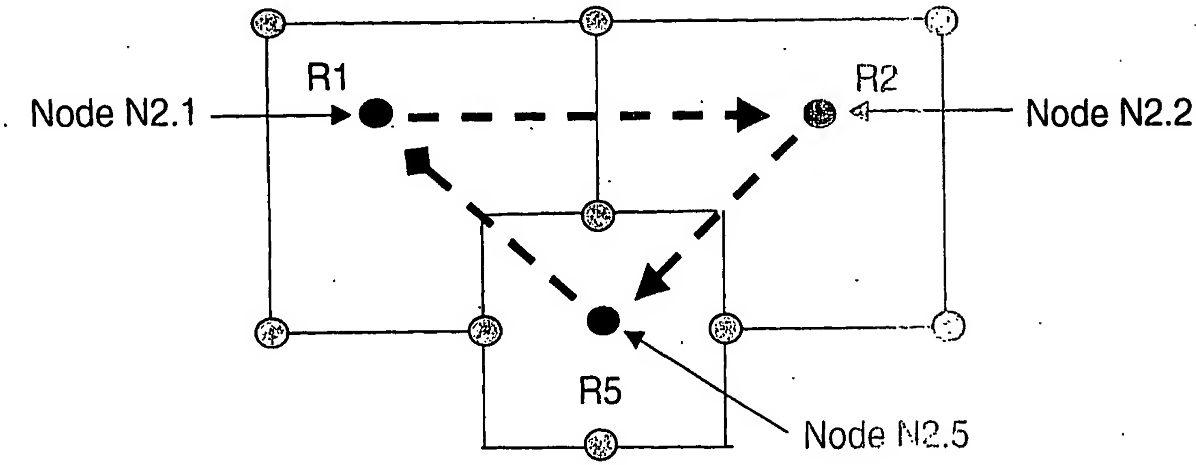


Figure 7

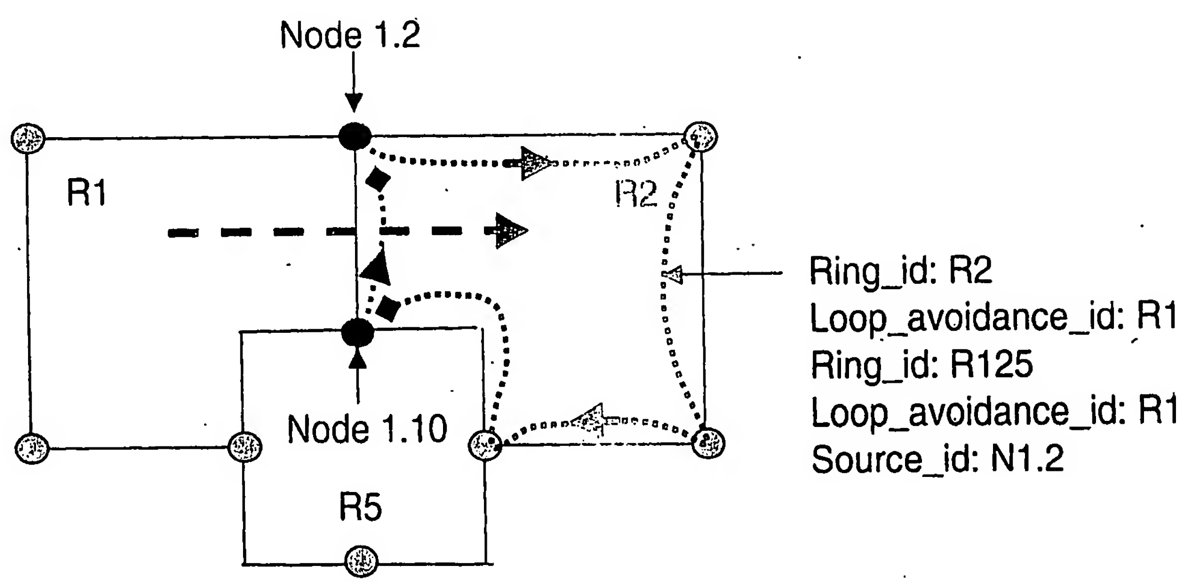


Figure 8

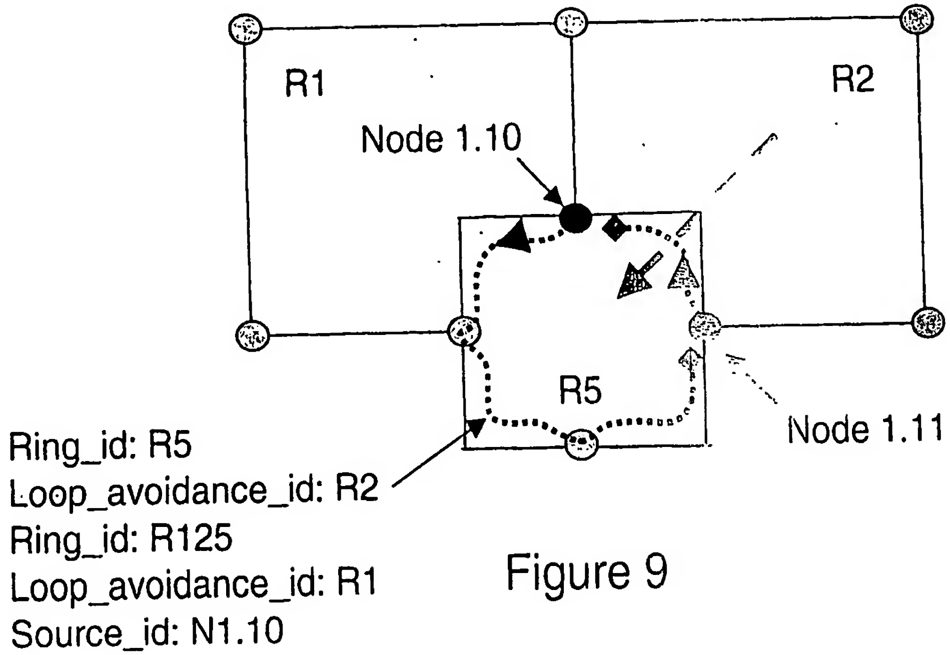


Figure 9

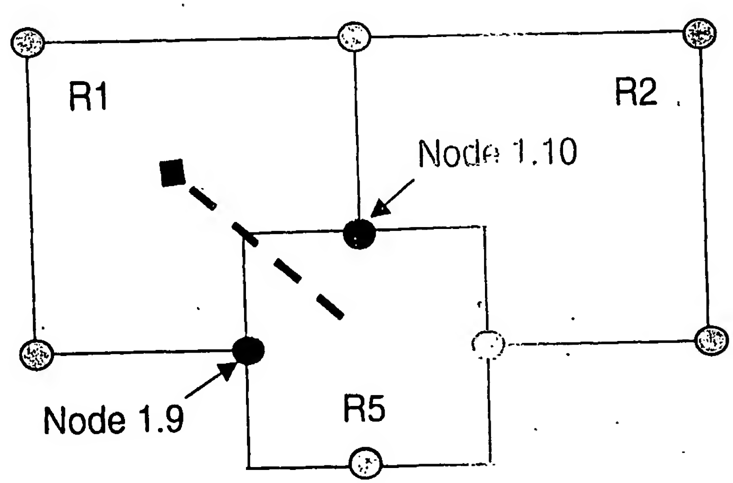


Figure 10

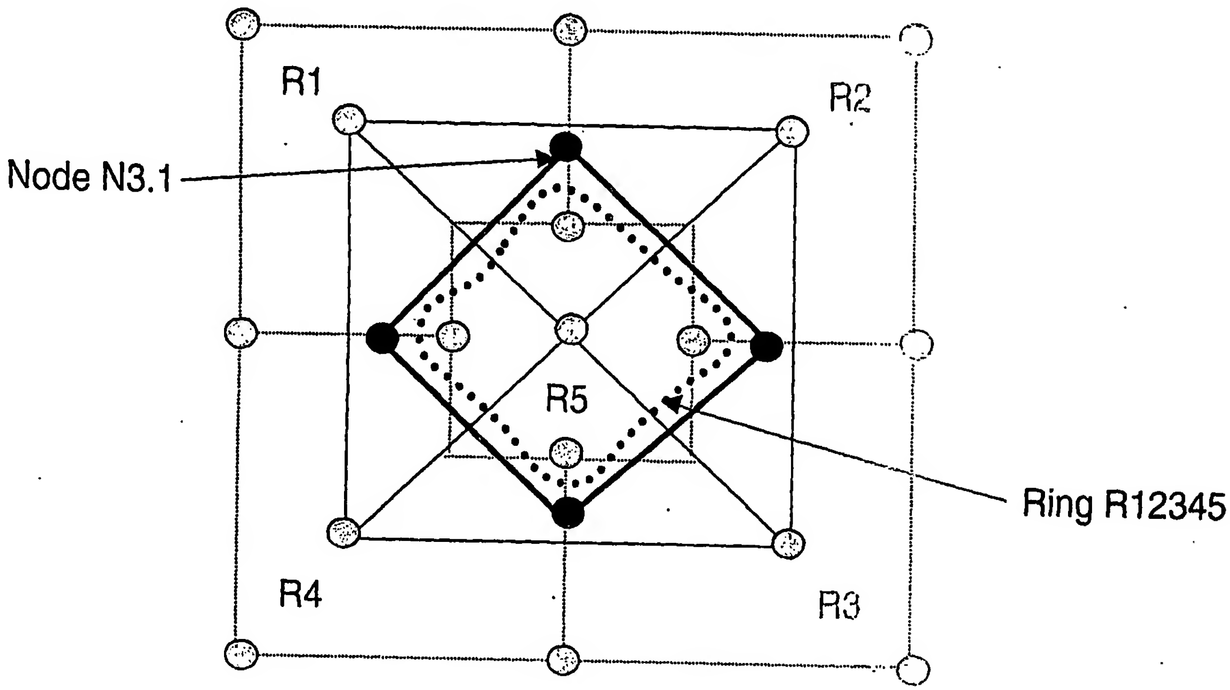


Figure 11

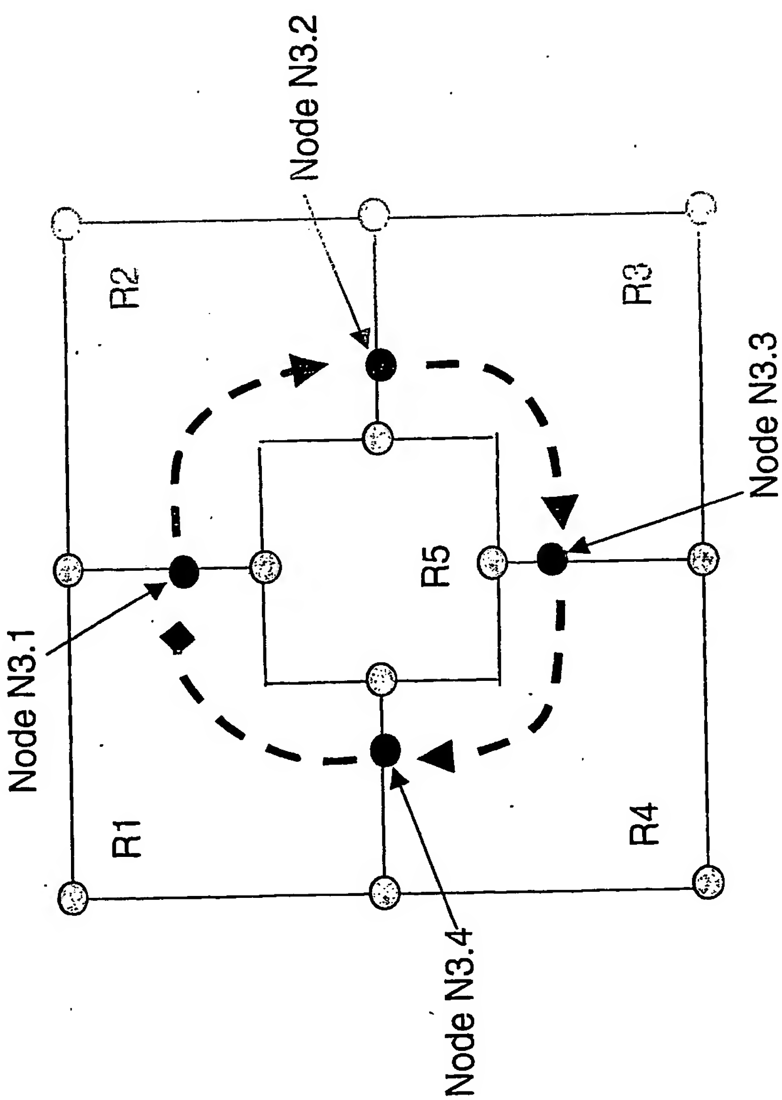


Figure 12

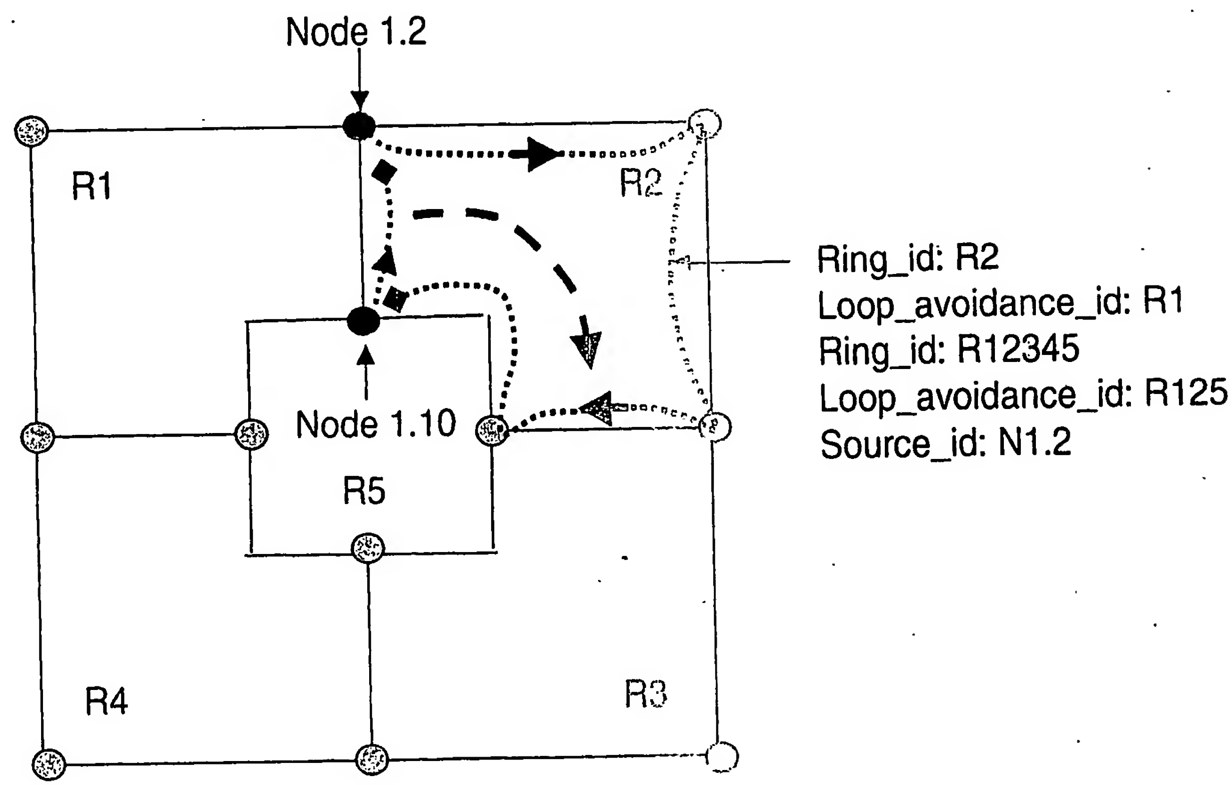

Node 1.2
R1
R2
Ring_id: R2
Loop_avoidance_id: R1
Ring_id: R12345
Loop_avoidance_id: R125
Source_id: N1.2
Node 1.10
R5
R4
R3

Figure 13

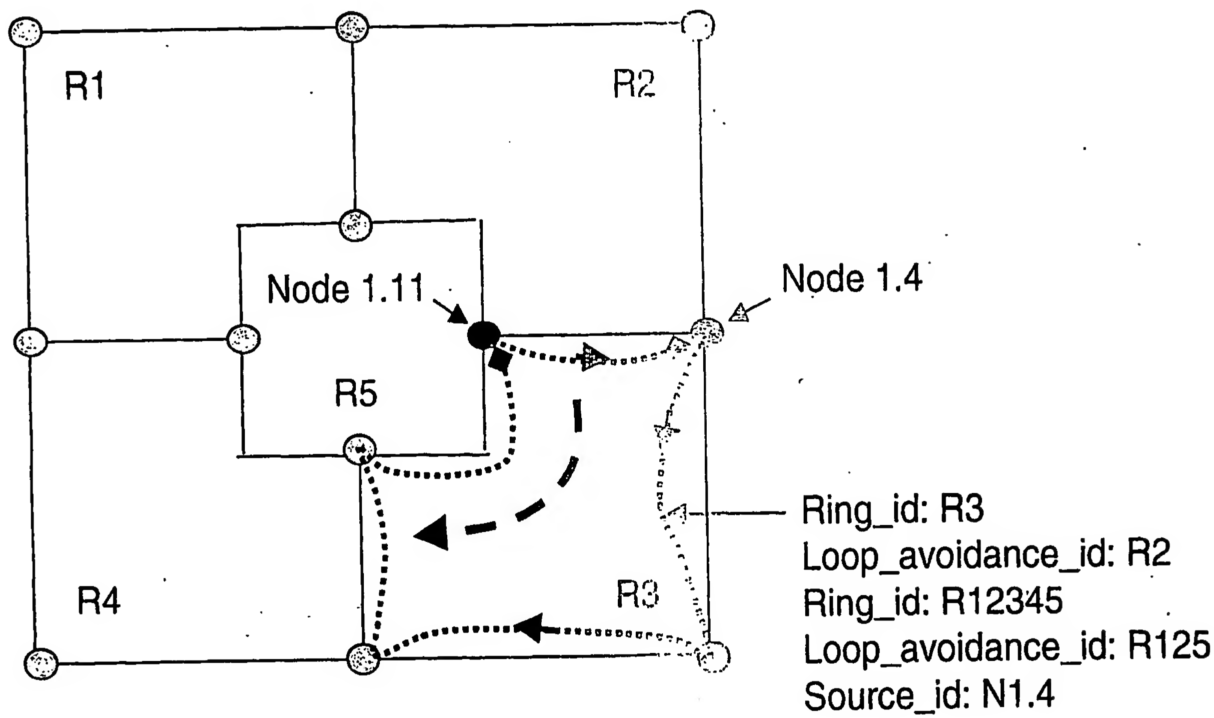
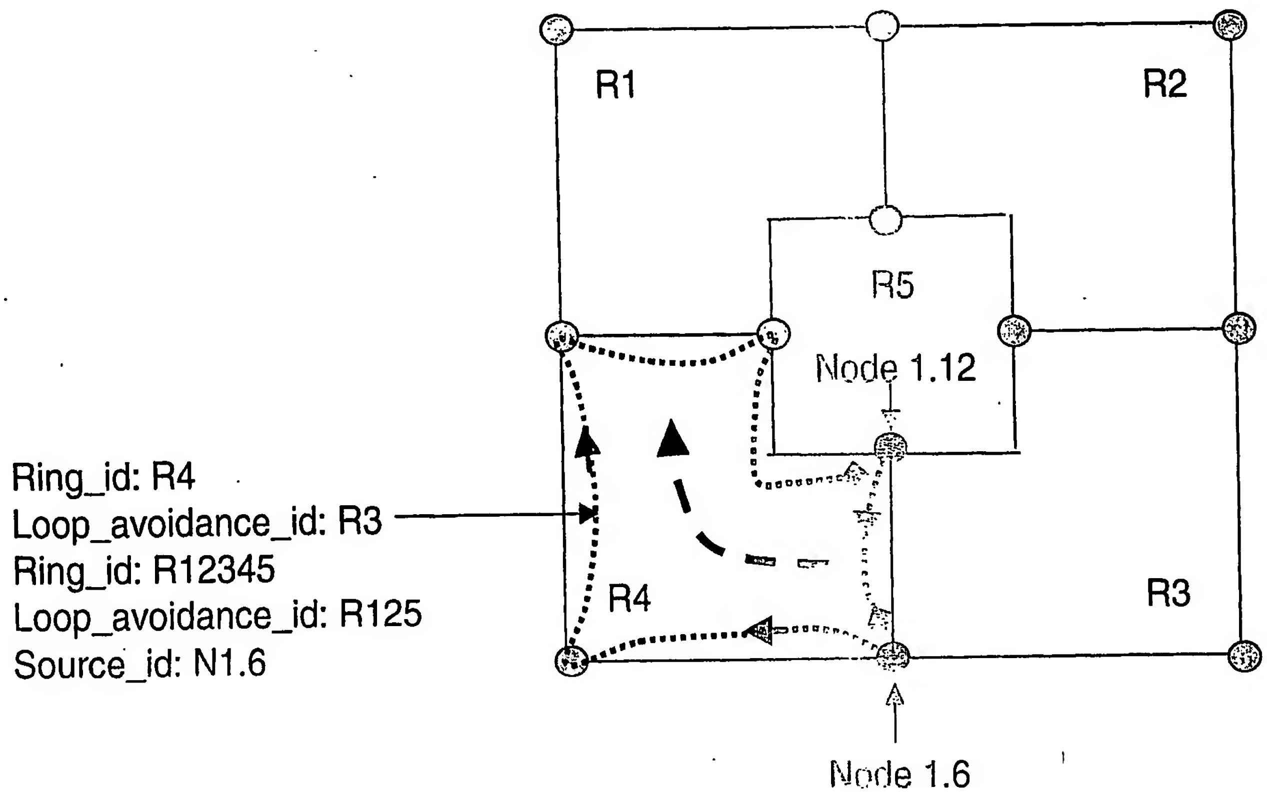

R1
R2
Node 1.11
Node 1.4
R5
R4
R3
Ring_id: R3
Loop_avoidance_id: R2
Ring_id: R12345
Loop_avoidance_id: R125
Source_id: N1.4

Figure 14



R1
R2
R5
Node 1.12
Ring_id: R4
Loop_avoidance_id: R3
Ring_id: R12345
Loop_avoidance_id: R125
Source_id: N1.6
R4
R3
Node 1.6

Figure 15

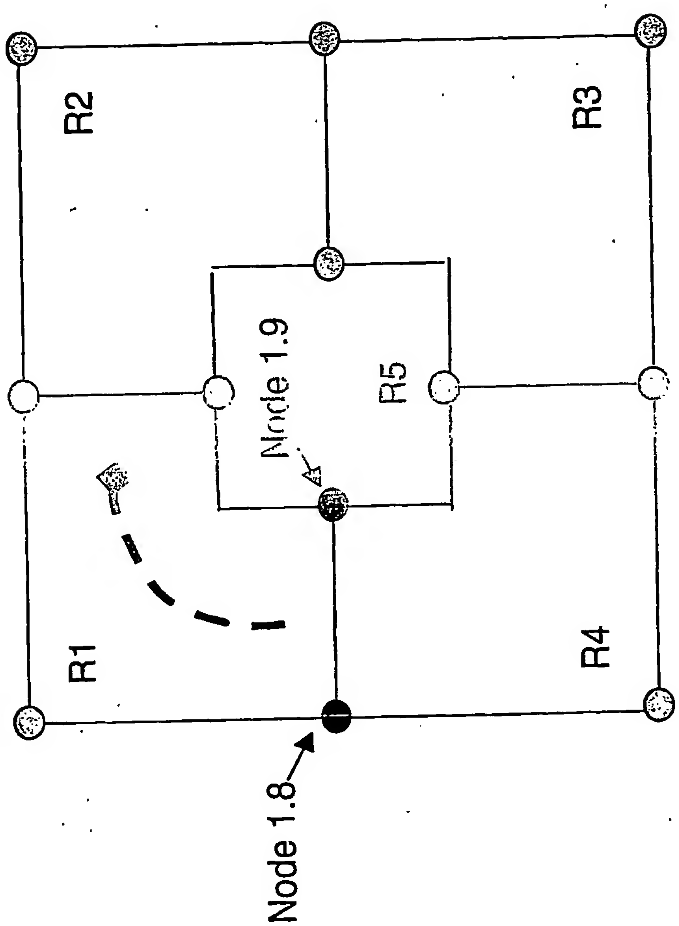


Figure 16

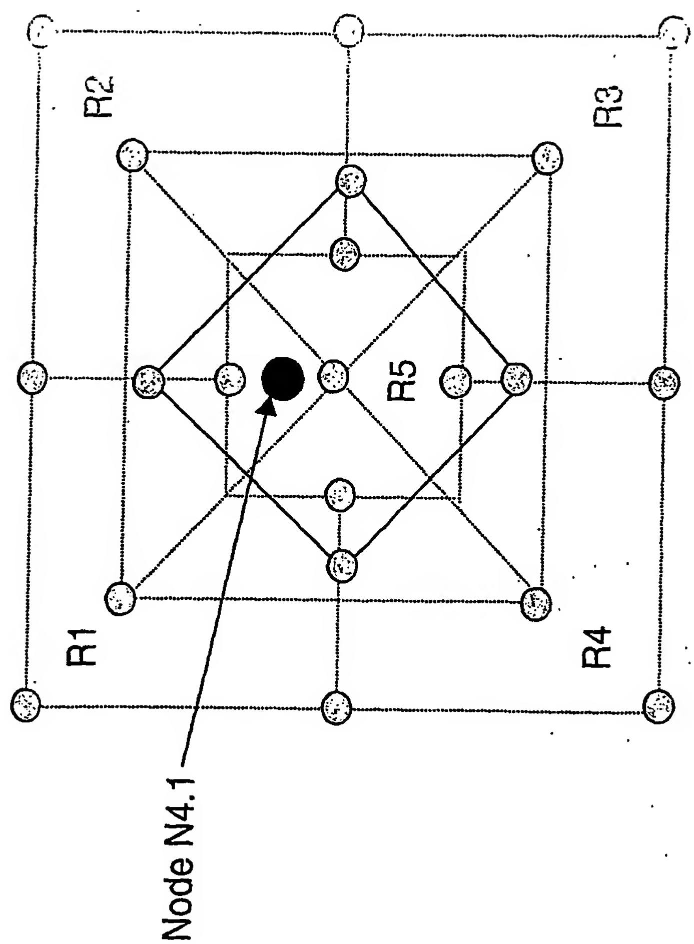


Figure 17

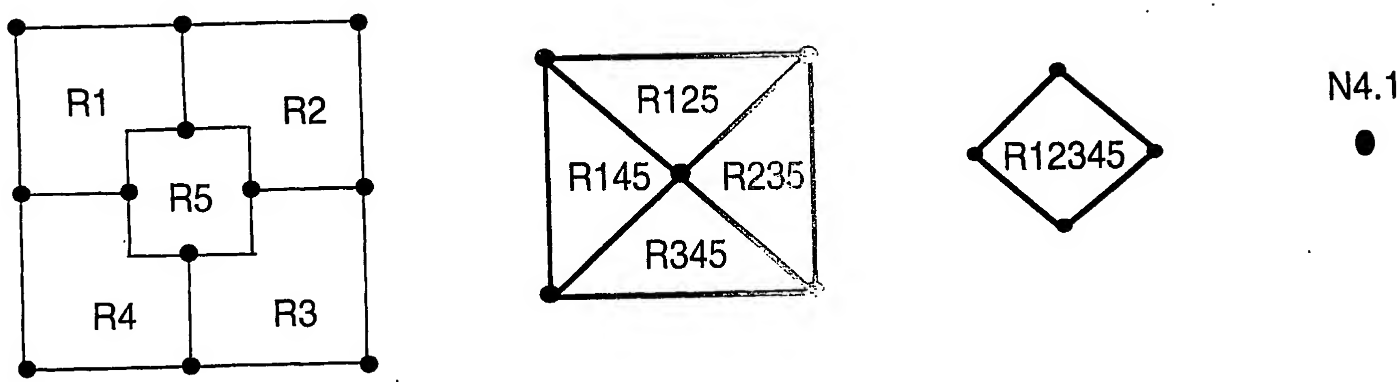


Figure 18

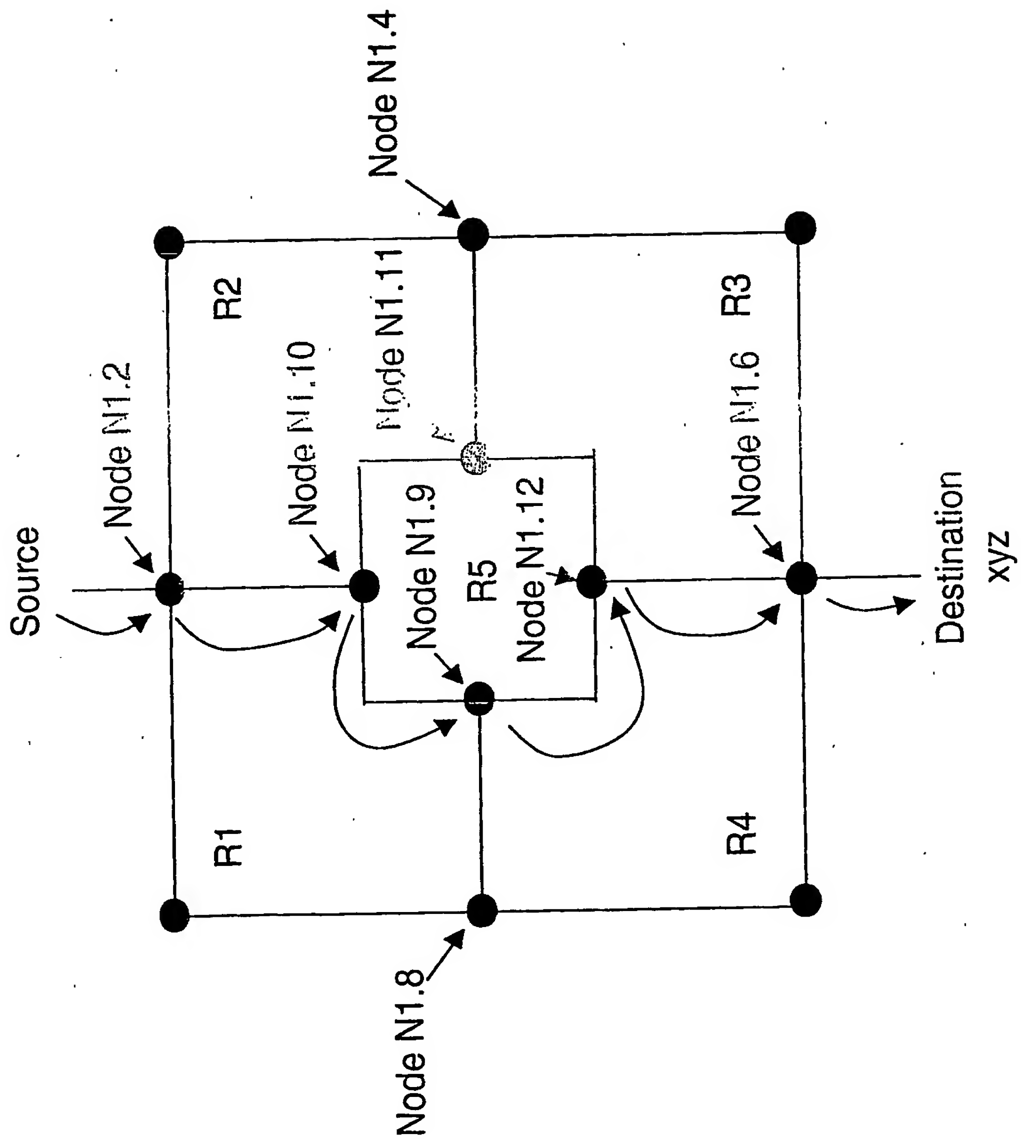

Source
Node N1.2
Node N1.10
Node N1.11
Node N1.4
R2
R3
Node N1.9
R5
Node N1.12
Node N1.6
Destination
xyz
R1
R4
Node N1.8

Figure 19

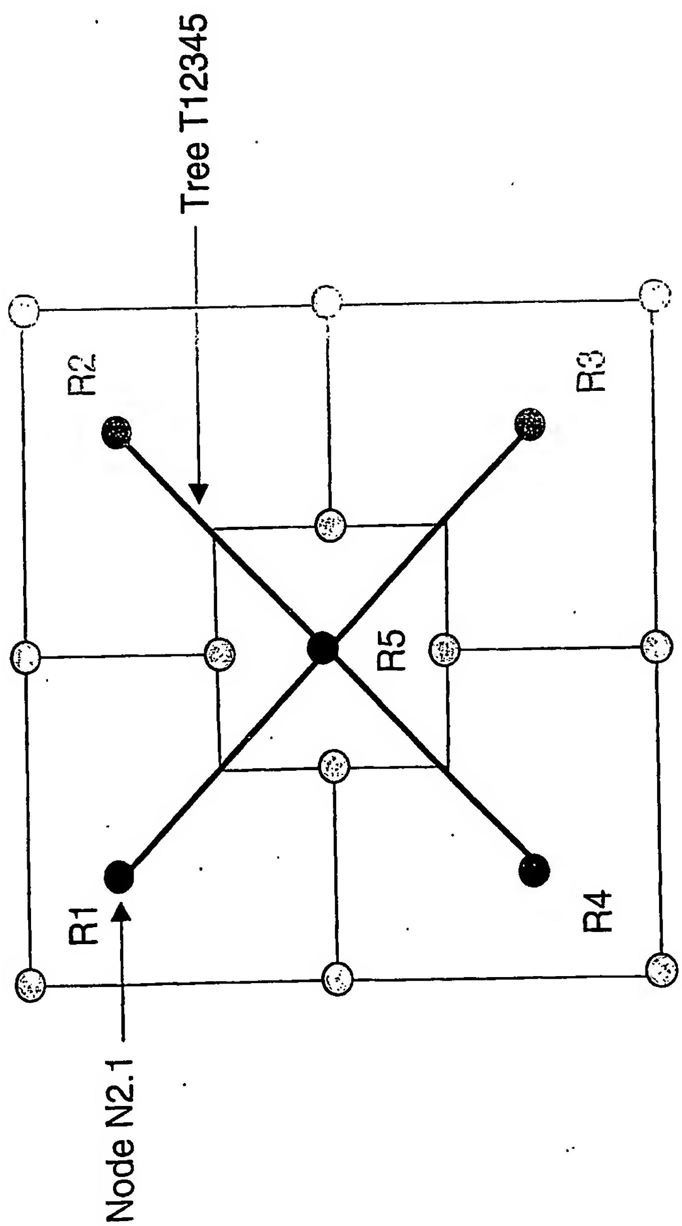


Figure 20

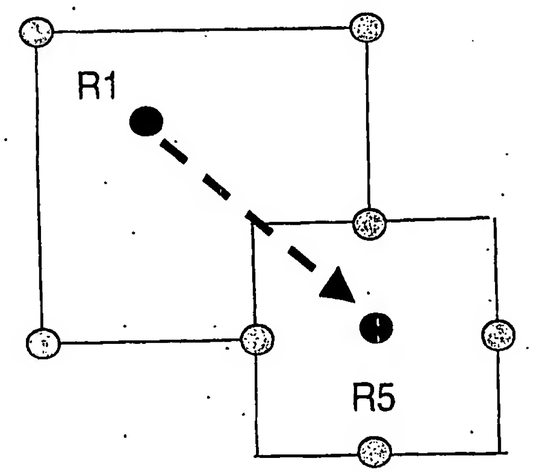


Figure 21

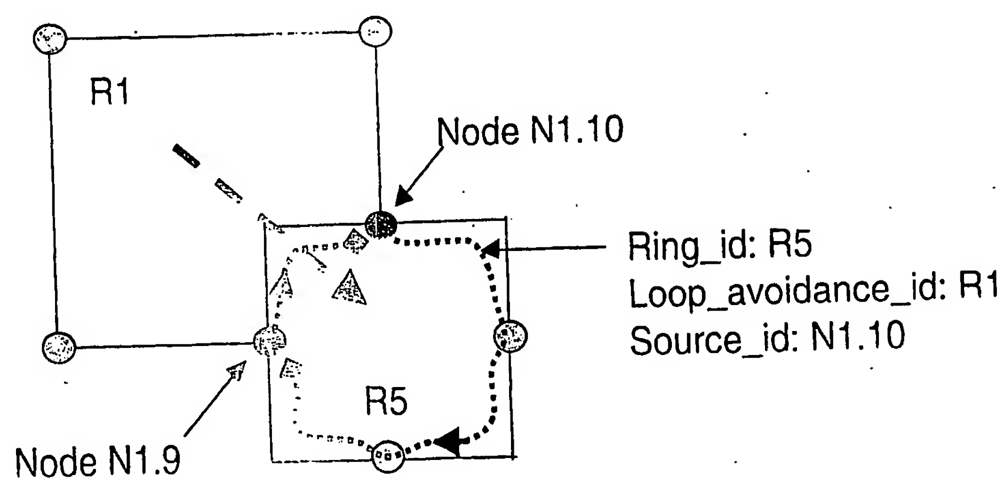


Figure 22

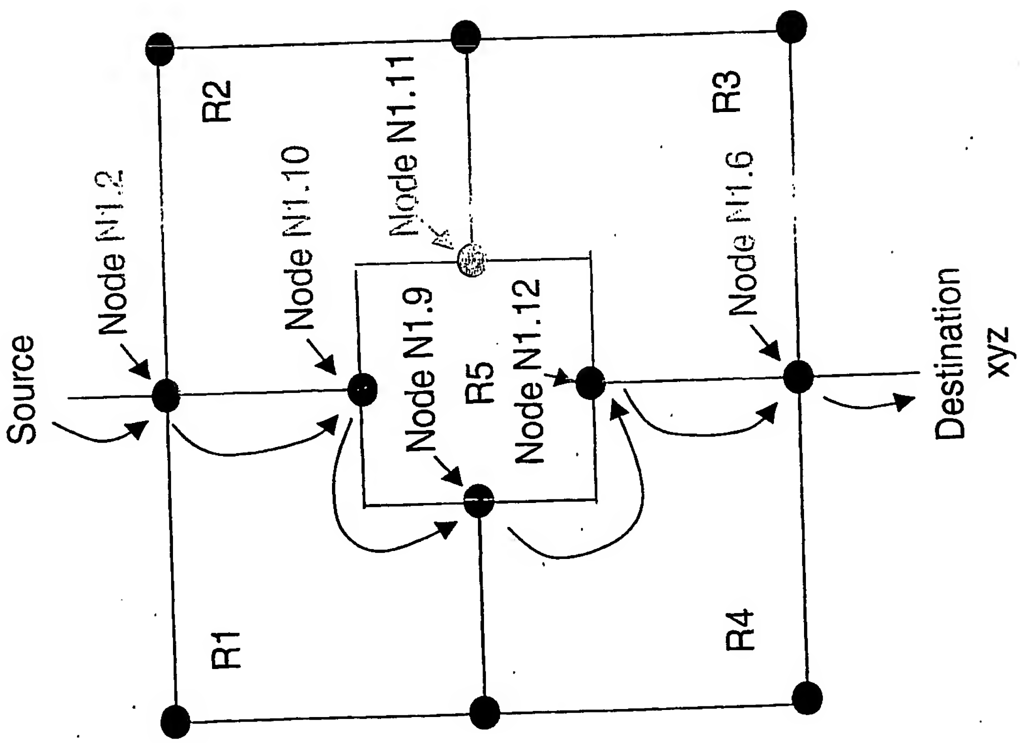

Source
Node N1.2
Node N1.10
Node N1.11
Node N1.9
R5
Node N1.12
Node N1.6
Destination
xyz
R1
R2
R3
R4

Figure 23

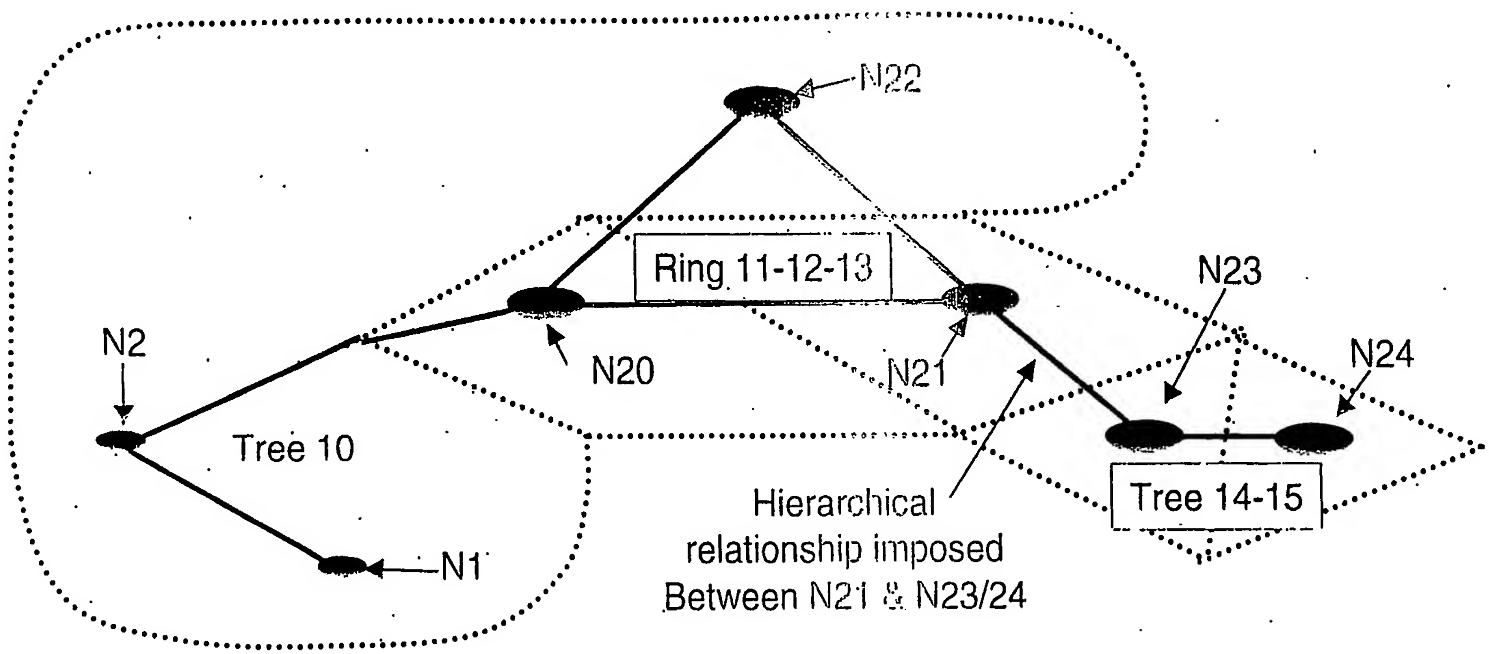


Figure 24

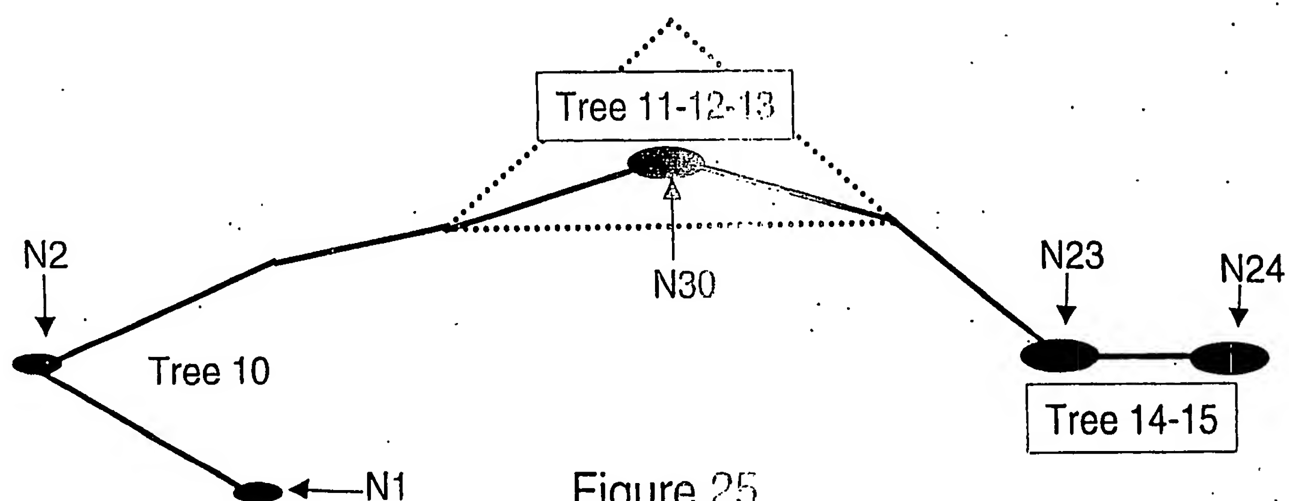


Figure 25